Elyx Member Journey Analytics – Technical Report

🎯 The Big Picture: What This App Accomplishes

Managing a client's long-term health and wellness journey often involves sifting through months of chat logs to track progress, frustrations, and key decisions. This is tedious, error-prone, and inefficient.

Elyx Member Journey Analytics automates this process. By uploading a single chat file, the system transforms raw text into a structured, interactive dashboard.

Key capabilities include:

1. **Conversation Decoding – Parses chats to identify participants, timestamps, and content.**
2. **Journey Structuring – Organizes the history into biweekly "Episodes" for clarity.**
3. **Insight Extraction – Captures metrics, topics, and emotional states per episode.**
4. **AI Summarization – Provides concise summaries of each two-week period.**
5. **Progress Visualization – Displays trends in sentiment, engagement, and topics through interactive charts.**
6. **AI Assistant – Enables query-based exploration using Retrieval-Augmented Generation (RAG).**

⚙️ Core Functions: The Engine Under the Hood

1. Data Handling & Structuring

- **parse_conversation_data(uploaded_file) → Reads raw logs, extracts sender, timestamp, and message.**
- **create_biweekly_episodes(conversations) → Segments conversations into 14-day "episodes," ensuring a continuous timeline.**

2. Episode Analysis Engine

- **analyze_episode(...) → Coordinates per-episode analysis (participants, topics, sentiment, metrics).**
- **detect_topics(conversations) → Identifies recurring themes such as *diet*, *workout*, *frustration*.**
- **calculate_metrics(conversations) → Computes message counts, ratios, and average team response times.**
- **estimate_consultation_time(sender, content) → Approximates consultation effort using heuristic rules.**

- **analyze_persona_states(...) & get_sentiment_from_openai(...) → Provides nuanced sentiment analysis via OpenAI.**
- **generate_episode_summary(conversations) → Uses GPT-4o-mini to create bullet-point summaries.**

3. AI Chat Assistant (RAG System) 🤠

- **initialize_vectorizer(conversations) → Converts conversation text into TF-IDF vectors.**
- **transform_query(question) → Expands queries with additional context via LLMs.**
- **retrieve_relevant_context(question) → Retrieves and re-ranks the most relevant passages using TF-IDF + Cross-Encoder.**
- **get_chatbot_response(question) → Generates context-grounded answers with GPT-4o-mini.**

4. User Interface & Visualizations 📊

- **main() → Manages Streamlit app layout and navigation.**
- **display_episode(...) → Presents structured episode cards with summaries, metrics, and states.**
- **display_chat_ui() → Builds the Q&A chat interface.**
- **Visualization Functions → Generate interactive topic charts, Sankey diagrams, and sentiment timelines using Plotly.**

🛠️ The Development Journey

Iterative UI and Model Development

- **Initial Version: Focused solely on summarizing the entire chat.**
- **Early Metrics: Randomized outputs tested with OpenRouter APIs (Gemma, Mistral).**
- **Optimization: Transitioned to smaller, faster paid API models for improved responsiveness.**
- **Prototype Chatbot: Introduced RAG for explainable, fact-grounded answers.**

Note: Current chatbot does not maintain memory; each query is independent.

Synthetic Data Generation

Creating realistic test data required multiple iterations:

- **Generation: Initial datasets produced with Gemini Pro using persona-driven prompts.**
- **Evaluation & Refinement: Improved using Perplexity for feedback and ChatGPT-4 for polish.**

- **Enhancement: DeepSeek was used to enrich persona traits and decision rationale.**
- **Humanization: Manual adjustments added natural quirks (family incidents, frustrations, anecdotes) for authenticity.**

The final dataset combines Gemini, ChatGPT, DeepSeek, Perplexity, and manual refinement, ensuring both technical accuracy and human realism.

🚀 Why It Matters

Elyx Member Journey Analytics provides:

- **Traceability – Every decision is explainable.**
- **Efficiency – Automates a process that otherwise takes hours.**
- **Insight – Highlights sentiment shifts, engagement levels, and evolving topics.**
- **Transparency – Builds trust through AI explanations grounded in actual data.**